Predicting of the Fibrous Filters Efficiency for the Removal Particles from Gas Stream by Artificial Neural Network

Érica Regina Filletti¹, Juliana Maria da Silva², Valdemir Garcia Ferreira³*

¹Department of Physical Chemistry, Institute of Chemistry, São Paulo State University-UNESP, Araraquara, Brazil
²Institute of Science and Technology, Federal University of Alfenas, Poços de Caldas, Brazil
³Department of Applied Mathematics and Statistics, Institute of Mathematics and Computer Sciences, University of São Paulo, São Carlos, Brazil
Email: erica@iq.unesp.br, juliana.silva@unifal-mg.edu.br

Received 4 May 2015; accepted 5 July 2015; published 8 July 2015

Abstract

In this paper, artificial neural networks are used for predicting single fiber efficiency in the process of removing smaller particles from gas stream by fiber filters. For this, numerical simulations are obtained of a classic model of literature for fiber efficiency, which is numerically solved along with the convection diffusion equation in polar coordinates for particle concentration, with associated initial and boundary conditions. A sufficient number of examples from two numerical simulations are employed to construct a database, from which parameters of a novel neural model are adjusted. This model is constructed based on the back propagation algorithm in order to map two features, namely Peclet number and packing density, which are extracted from the numerical simulations into the corresponding single fiber efficiency. The results indicate that the developed neural model can be trained in a reasonable computational time and is capable of estimating single fiber efficiency from examples of the test set with a maximum error of 1.7%.

Keywords

Artificial Neural Network, Back Propagation Algorithm, Fiber Filters, Particle Capture

1. Introduction

Concern with the increasing air pollution has stimulated the development of new techniques that operate in the

efficiency of solid-gas separation, mainly for smaller particles. Filtration by fibrous filters is one of the most widely used techniques to capture fine particles from a gas stream. The process takes place by passing the aerosol through the filter media, so that the particles will be deposited on the filter surface [1]. A fibrous filter consists of many threadlike fibers, of varying sizes, positioned more or less normal to the direction of the fluid flow. The fluid passes through the regions between the fibers and particles, which are suspended in the fluid, are removed by their collision and attachment to the fiber surface. The fibrous filters are notable for their usefulness in removing fine particles from aerosols at low cost and easy implementation [2].


Thus, it is evident from the works above that theoretical studies are of prime importance for understanding the performance of fibrous filters. Also, it is important to observe that, combined with artificial neural network (ANN), predictions from parameters of this process may become an important factor for innovation and development of new techniques. In addition, ANNs have several advantages that make them attractive tools: they can be trained to learn from examples; are easily updated and generalized; possess a large degree of freedom and accurate prediction at high speed [12]-[14]. ANNs possess the ability to identify a relationship from given patterns and this makes it possible to solve large-scale complex problems such as nonlinear mathematical models.

Recently, ANNs models have been successfully applied to various branches of science and technology, such as medicine [15] [16], mathematics and engineering [17], among others. However, these references failed to mention that ANNs have been applied to predicting the collection efficiency of fibrous filters; this being the prime motivation for the present study.

The objective of this work is to investigate the possibility of using a neural network to directly assess the single fiber efficiency. Thus, a much more robust measurement method can be developed. For this, in order to construct a database for developing a neural model, a mathematical model is implemented and numerical results are obtained for several Peclet numbers and packing density, which are important parameters in this process.

2. Modelling for Single Fiber Efficiency

This section aims to present the mathematical modelling to simulate the process of particle deposition on a single fiber to construct a database for an ANN model.

As the structure of fibrous filters is complex, many works have adopted cell models in which fluid flow equations are solved in a two dimensional cell surrounding the single fiber, where the boundary conditions take into account neighboring fibers [2]. Thus, the cell model was adopted to determine the flow field. The geometry is illustrated in Figure 1, where \( a = 1 \) is the dimensionless fiber radius, \( b = a^{\frac{1}{2}} \) is the dimensionless cell radius that is determined by the filter packing density \( \alpha \) and \( U \) is the flow velocity scale [7] [18] [19].

2.1. The Physical Model for Single Fiber Efficiency

In this study, the deposition of aerosol particles in fibrous filters is considered to occur exclusively because of diffusion, which is one of the major particle collection mechanisms, especially when dealing with smaller particles (see [2] [3] [18]).

The single fiber efficiency is mathematically expressed by

\[
\eta = \frac{2}{Pe} \int_0^\infty \tilde{c}_n \frac{\partial r}{\partial r} d\theta \bigg|_{-1},
\]
where \( n \) is the particle concentration and \( Pe \) is the Peclet number.

To determine numerically the value of the integral in Equation (1), the following dimensionless convection diffusion partial differential equation is utilized, which in polar coordinates \( r \) and \( \theta \) is given by [3]

\[
\frac{u_r}{r} \frac{\partial n}{\partial r} + \frac{u_\theta}{r} \frac{\partial n}{\partial \theta} = \frac{2}{Pe} \left( \frac{\partial^2 n}{\partial r^2} + \frac{1}{r} \frac{\partial n}{\partial r} + \frac{1}{r^2} \frac{\partial^2 n}{\partial \theta^2} \right),
\]

(2)

where \( u_r \) and \( u_\theta \) are, respectively, the radial and angular components of the gas velocity, expressed as

\[
u_r = \frac{1}{r} \frac{\partial \psi}{\partial \theta} \text{ and } u_\theta = - \frac{\partial \psi}{\partial r},
\]

(3)

with \( \psi \) being the dimensionless Kuwabara function [5] [18].

For fibrous filtration, the third term on the right-hand side of Equation (2) is much smaller than the other terms and can be neglected to give [2] [3]

\[
\frac{u_r}{r} \frac{\partial n}{\partial r} + \frac{u_\theta}{r} \frac{\partial n}{\partial \theta} = \frac{2}{Pe} \left( \frac{\partial^2 n}{\partial r^2} + \frac{1}{r} \frac{\partial n}{\partial r} \right).
\]

(4)

The adopted boundary conditions are

\[
\begin{align*}
    r &= 1, \ 0 \leq \theta \leq \pi, \ n = 0, \\
    r &= b, \ 0 \leq \theta \leq \pi, \ n = 1, \\
    1 \leq r \leq b, \ \theta = 0, \ \theta = \pi, \ \frac{\partial n}{\partial \theta}(r, \theta) &= 0
\end{align*}
\]

(5)-(7)

2.2. Numerical Procedure

Equation (4) associated with the boundary conditions (5)-(7) was discretized by finite-difference methodology on a staggered grid system; the diffusive terms were approximate by second order central differences and the non-linear convection terms were approximated by the TOPUS upwind scheme [20], which was expressed in this work as

\[
\begin{align*}
    n_f &= \begin{cases} 
        n_{U_f} + \left(n_{D_f} - n_{R_f}\right) \left(\beta \hat{n}_{U_f} + c_1 \hat{n}_{U_f} + c_2 \hat{n}_{U_f} + c_3 \hat{n}_{U_f}\right) & \hat{n}_{U_f} \in (0,1) \\
        n_f = n_{U_f}, & \hat{n}_{U_f} \notin (0,1)
    \end{cases} \\
    \hat{n}_{U_f} &= \left(n_{D_f} - n_{R_f}\right) / \left(n_{D_f} - n_{R_f}\right)
\end{align*}
\]

(8)

where \( f \) is a boundary face of a computational cell (see Figure 2). The quantity \( \hat{n}_{U_f} \) corresponds to the normalized variables of Leonard [21]. The coefficients \( c_1, c_2 \) and \( c_3 \) are expressed as \( c_1 = -2\beta + 1, \ c_2 = (5\beta - 10)/4 \) and \( c_3 = (-\beta + 10)/4 \), where \( \beta \) is an adjustable real parameter chosen in the interval \([-2, 2]\) to attend the total variation diminishing (TVD) condition of Harten [22]. The neighboring nodes \( D_f \) (Downstream), \( U_f \) (Upstream) and \( R_f \) (Remoteupstream) of a computational face \( f \) are prescribed according to the advection velocity \( \nu_f \) (\( u_r \) or \( u_\theta \)) at this face, as can be observed in Figure 2. The TOPUS scheme was chosen for
this study since the use of conventional schemes (central differences and first order upwind) has been unsatisfactory for predicting single fiber efficiency in the range $100 \leq Pe \leq 165$, $0.2 \leq \alpha \leq 0.3$.

The linear system resulting from the discretization of Equation (4) was solved by the Gauss-Seidel iteration procedure. The numerical results were compared favorably with experimental data of literature and published in [7] [19].

3. Artificial Neural Network

In this section, the basic concepts of ANNs, the back propagation algorithm and a new neural model for predicting $\eta$, in the process of removing aerosol particles, are introduced.

3.1. Theoretical Basis

In short, ANNs are a form of artificial intelligence composed of a network of connected nodes (neurons). Formally, an ANN can be defined as nonlinear mapping of an input onto an output vector space. This is achieved through layers of neurons in which the input coordinates are summed according to specific weights and biases to produce a single output. Finding suitable architecture for ANNs is extremely critical; most works in the literature use trial and error to adjust the parameters of a neural network (see, for example, [12] [23]-[25].)

In this study, a four-layer feed-forward ANN with two hidden layers was employed and the back propagation algorithm was used for the training. No recursiveness was adopted here, that is, the input vector of a specific neuron layer was formed only by the values of the preceding layer. A schematic representation of a typical ANN structure is depicted in Figure 3.

The notation in a feed-forward network is as follows. If the activation function of $i$-th neuron in the $j$-th layer is indicated by $F_{i,j}()$, its output $s_{i,j}$ can be calculated from outputs of the preceding layer $s_{i,j-1}$ and from the corresponding biases $b_{i,j}$ and weights $w_{k,i,j}$ (the second subscript $k$ indicates the neuron in the $(j-1)$-th layer from which the connection is being established), according to

$$s_{i,j} = F_{i,j} \left( b_{i,j} + \sum_{k} w_{k,i,j} s_{k,j-1} \right).$$

(9)

After denoting the input and output network values by $\sigma_i$, $i = 1, 2, 3, 4$ and $\xi_j$, $j = 1, 2$, respectively, the mapping relationship of one onto the other can then be done by applying Equation (9). For instance, the mapping in Figure 3 reads

$$\xi_j = F_{1,3} \left( b_{1,3} + \sum_{k=1}^{2} w_{k,1,2} F_{k,2} \left( b_{2,2} + \sum_{m=1}^{2} w_{k,m,1} F_{m,1} \left( b_{m,3} + \sum_{n=1}^{2} w_{m,n,0} \sigma_n \right) \right) \right) \right)$$

(10)

Equation (10) makes it clear that the relationship between $\sigma_i$ and $\xi_j$ is unambiguously defined by choosing the activation functions and by setting the biases and weights. One very important characteristic of ANNs, among others, is the so-called learning potential, that is, the possibility of adjusting the biases and weights through a convenient training rule to closely reproduce pre-assigned pairs of input/output values. Back propagation is probably the most used training heuristic and it is particularly well adapted to feed-forward architecture. It is based on the iterative application of a discrete gradient descent algorithm, computed from the first derivatives of a conveniently defined error function whose arguments are the parameters of the network (weights and biases). In general, the basic steps of the back propagation procedure, which is implemented in this work, are the following [26]:
1) Initialize the ANN’s parameters $b_{i,j}$ and $w_{i,k,j}$ with random numbers;
2) Take the $p$-th $(\sigma_i^p, \delta_i^p)$ pair from a training data set with pre-assigned input/output pairs;
3) Calculate the outputs with the same input and form the pair $(\sigma_i^p, \xi_i^p)$;
4) Evaluate the error $e$ between the desired and the obtained output values according to the $L_2$ norm

$$e = \sqrt{\sum (\delta_i^p - \xi_i^p)^2}$$

(11)

5) Calculate the derivatives of the error with respect to $b_{i,j}$ and $w_{i,k,j}$;
6) Modify the ANN parameters according to the steepest descent strategy and a specified learning rate $\gamma$

$$b_{i,j} \leftarrow b_{i,j} - \gamma \frac{\partial e}{\partial b_{i,j}}$$

(12)

$$w_{i,j} \leftarrow w_{i,k,j} - \gamma \frac{\partial e}{\partial w_{i,k,j}}$$

(13)

7) Iterate from 2 to 5, by successively modifying $b_{i,j}$ and $w_{i,k,j}$ until a defined number of learning epochs (cycles) or a convenient stopping criterion has been achieved.

The performance of an ANN is affected by internal architecture (number of hidden layers and number of neurons in each one) and type of interconnections (feed-forward, recursiveness, winner-take-all, etc.). The exact shape of the activation function has limited effects on the overall performance and is usually set according to the needs of the training heuristics (a sigmoid function in the case of the back propagation method). There is no general mathematical theory but rather a number of empirical rules that must be considered when constructing such models.

### 3.2. Neural Modelling for Single Fiber Efficiency

The ANN implemented in this work has two inputs and one output. Various architectures were trained and tested; the neural model that provided good results, obtained by trial and error, was that containing two intermediate layers with six and three neurons. The number of neurons in the intermediate layers was adjusted to enable the ANN to learn complex tasks for the gradual extraction of significant information from the inputs [27] [28]. The activation functions used in the neural network were the tangent sigmoid in the intermediate layers and a linear function in the output layer, expressed respectively as

$$\phi(x) = \frac{2}{(1 + \exp(-2x))} - 1$$

(14)

$$\phi(x) = x$$

(15)

where $v = b_{i,j} + \sum_k w_{i,k,j} s_{k,j-1}$ is the activation potential of the above functions.
The neuron of the output layer is responsible for estimating single fiber efficiency ($\eta$). The training procedure uses the acquisition of the Peclet number ($Pe$) and packing density ($\alpha$) for the filter to a sufficient level for inference of the $\eta$. Two numerical simulations were performed with $Pe$ and $\alpha$ varying uniformly in the ranges: a) $20 \leq Pe \leq 80$ and $0.1 \leq \alpha \leq 0.2$; and b) $100 \leq Pe \leq 165$ and $0.2 \leq \alpha \leq 0.3$. Equation (1) was used for estimating the $\eta$.

In both simulation cases, the input data matrix has 2 lines for 165 columns with the line number corresponding to the number of inputs ($Pe$ and $\alpha$) and the column to the examples used in the training of the ANN; each one corresponds to a known and fixed value of the $\eta$. The output matrix has 1 line for 165 columns, where 1 is the number of desired output and 165 is the number of examples. Generalization was considered: 88 different examples, from those used in the training process, were presented to the ANN in the first simulation; and 66 examples in the second simulation. Therefore, in the first calculation, 253 pairs of input/output values were obtained to construct the database, and in the second calculation 231 pairs were employed. The division of the examples in the training and test sets was done randomly and some divisions were tested. The ANN results for all the tested divisions were shown to be practically equal, with the same correlation coefficients for the data.

4. Results and Discussion

The purpose of this section is that it presents the training error of neural network developed and the results obtained for the data of the test.

The results of the training of the ANN have shown that it is capable of reproducing the input/output relationship of the training set data. To evaluate the generalization capacity of the ANN, the characteristic data of new numerical tests were presented to the neural networks and a good correlation between the input and output data was observed.

In attempting to reduce the error in a reasonable time, optimization of the ANN parameters (learning rate, epochs and number of neurons in the intermediate layers) was performed. The learning rate used in both tests was 0.1 and $12\times10^4$ epochs were considered, with a training time of 36 minutes for the first case, and 44 minutes for the second one. As can be observed from Figure 4 and Figure 5, the training error for both ANNs behaved as expected like $O(10^{-4})$.

Figure 6 and Figure 7 present the estimated data for $\eta$, calculated by the ANN after the training period, considering the examples in the test set of the first and second range. The dots in the graphs are the values determined by the ANN. For the first range (see Figure 6), the straight line that best represents the correlation between the values of $\eta$ estimated by the ANN and the numerical results is $y = 1.0047x - 0.0012 \ [29] \ [30]$. The Pearson $r$-squared correlation coefficient ($R^2$) between the numerical values and the ANN values was 0.9996, showing that there was a strong relationship between the input and output data provided by the ANN; whereas, in the second range, the best representation for $\eta$ is $y = 1.0021x - 0.0003$ with a correlation coefficient of 0.9995. These results indicate that training of the ANNs was successfully accomplished.
In summary, the results obtained through the ANN for the test set examples were quite satisfactory, with maximum relative errors of 1.7% for the first case and 0.9% for the second case. Table 1 presents (for some test set examples) the responses of the ANN for $\eta$ in the two ranges, where good agreement can be seen between the numerical $\eta$ ($\eta_{\text{Numeric}}$) and the $\eta$ estimated by the ANN ($\eta_{\text{ANN}}$).
5. Conclusions

Artificial neural models were trained with data derived from numerical simulations in order to determine smaller particle capture efficiency in a single fiber. To construct a database, a sufficient number of examples were performed, from which the ANN parameters (weights and biases) were adjusted. A model was developed to map the parameters $Pe$ and $\alpha$ with the aim of estimating the single fiber efficiency, $\eta$. The developed ANN was shown to be capable of estimating correct values for fiber efficiency from examples not contemplated in the training.

From this research, it can be inferred that, in the context of fiber filter efficiency, the developed neural model is an effective tool for performing nonlinear mapping of the parameters. The simplicity and efficiency of the proposed neural approach indicates that the present methodology can be effectively used in the process studied. Furthermore, it should be emphasized still that, although numerical data have been used in this work, the results obtained suggest that if the experimental data are available, these data could have been used to train the neural network which would produce a very powerful prediction tool.

Acknowledgements

The authors wish to express their thanks to the Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq-Grants No. 475390/2011-3) for the financial support.
References


## Nomenclature

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>$a$</td>
<td>fiber radius dimensionless</td>
</tr>
<tr>
<td>$b$</td>
<td>cell radius dimensionless</td>
</tr>
<tr>
<td>$b_{i,j}$</td>
<td>bias of $i$-th neuron in the $j$-th layer</td>
</tr>
<tr>
<td>$D$</td>
<td>Brownian diffusion coefficient</td>
</tr>
<tr>
<td>$e$</td>
<td>error between the desired and the obtained output</td>
</tr>
<tr>
<td>$F_{i,j}()$</td>
<td>activation function of $i$-th neuron in the $j$-th layer</td>
</tr>
<tr>
<td>$n$</td>
<td>particle concentration</td>
</tr>
<tr>
<td>$n_{Df}$</td>
<td>particle concentration at location $D_f$</td>
</tr>
<tr>
<td>$n_{Rf}$</td>
<td>particle concentration at location $R_f$</td>
</tr>
<tr>
<td>$n_{Uf}$</td>
<td>particle concentration at location $U_f$</td>
</tr>
<tr>
<td>$Pe$</td>
<td>Peclet number ($= 2aU/D$)</td>
</tr>
<tr>
<td>$R$</td>
<td>polar coordinate dimensionless</td>
</tr>
<tr>
<td>$R^2$</td>
<td>correlation coefficient</td>
</tr>
<tr>
<td>$s_{i,j-1}$</td>
<td>output of $i$-th neuron in the $(j-1)$-th layer</td>
</tr>
<tr>
<td>$s_{i,j}$</td>
<td>output of $i$-th neuron in the $j$-th layer</td>
</tr>
<tr>
<td>$U$</td>
<td>flow velocity scale</td>
</tr>
<tr>
<td>$u_r$</td>
<td>radial gas velocity</td>
</tr>
<tr>
<td>$u_\theta$</td>
<td>angular gas velocity</td>
</tr>
<tr>
<td>$v_f$</td>
<td>advection velocity</td>
</tr>
<tr>
<td>$v$</td>
<td>activation potential</td>
</tr>
<tr>
<td>$w_{i,j,k-1}$</td>
<td>weights of $i$-th neuron in the $(k-1)$-th layer</td>
</tr>
<tr>
<td>$w_{i,j,k}$</td>
<td>weights of $i$-th neuron in the $k$-th layer</td>
</tr>
</tbody>
</table>

- $(w_{i,j,k})$ indicates the neuron in the $(j-1)$-th layer from which the connection is being established.
- $(w_{i,j,k})$ indicates the neuron in the $k$-th layer from which the connection is being established.

## Greek

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\alpha$</td>
<td>filter packing density</td>
</tr>
<tr>
<td>$\beta$</td>
<td>adjustable parameter</td>
</tr>
<tr>
<td>$\gamma$</td>
<td>learning rate</td>
</tr>
<tr>
<td>$\delta_i$</td>
<td>training data output values</td>
</tr>
<tr>
<td>$\eta$</td>
<td>single fiber efficiency</td>
</tr>
<tr>
<td>$\eta_{ANN}$</td>
<td>single fiber efficiency estimated by ANN</td>
</tr>
<tr>
<td>$\eta_{Num}$</td>
<td>numerical single fiber efficiency</td>
</tr>
<tr>
<td>$\theta$</td>
<td>polar coordinate dimensionless</td>
</tr>
<tr>
<td>$\zeta_i$</td>
<td>network output values</td>
</tr>
<tr>
<td>$\sigma_i$</td>
<td>network input values</td>
</tr>
<tr>
<td>$(\sigma^p, \varphi^p)$</td>
<td>outputs of the network with the same input</td>
</tr>
<tr>
<td>$(\sigma^p, \delta^p)$</td>
<td>the $p$-th pair pre-assigned input/output</td>
</tr>
</tbody>
</table>